DEPARTMENT OF COMMERCE
CHARLES SAWYER, Secretary

WEATHER BUREAU
F. W. REICHELDERFER, Chief

MONTHLY WEATHER REVIEW

Editor, JAMES E. CASKEY, JR.

Volume 80
Number 7

JULY 1952

Closed September 15, 1952
Issued October 15, 1952

MEASURES OF THE VARIABILITY OF PRECIPITATION

RICHMOND W. LONGLEY

Meteorological Division, Department of Transport, Toronto, Canada

[Manuscript received June 9, 1952]

ABSTRACT

A study is made of the usefulness of four different measures of the relative variability of precipitation. It is discovered that the two measures, the difference between the extremes divided by the median, are subject to too great fluctuations in sampling to be satisfactory. There is little to choose between the two measures, the mean deviation or the standard deviation divided by the mean. Even with these measures, the fluctuations are sufficiently large so that a period of 30 years is desirable before computing them. It is further shown that the standard deviation divided by the mean varies slightly with mean precipitation but, in spite of this, it is a satisfactory measure with which to compare variability of precipitation in different localities.

INTRODUCTION

In considering precipitation and its importance in human affairs, the most significant information is the average amount that falls during the year. The average annual trend of precipitation is almost as important as the total rainfall. These statistics are commonly found in climatological tables.

A third important function of the precipitation is its variation from one year to the next. As an example of the value of knowing this function, consider a region where the demand for water almost equals the average supply. The need might be for the development of water power, or for the growth of a certain kind of crop. If there is little variation from year to year, the supply will seldom fail. If, instead, the variation from year to year is great, there will be some years when there is a surplus, while in other years the supply will not meet the demand. In spite of its value to mankind, a relatively small amount of work has been done on precipitation variability.

In statistical theory, there are three commonly used measures of variability, the standard deviation, the mean deviation, and the quartile deviation or the semi-interquartile range. The standard deviation, σ, is defined by the equation

$$\sigma=\sqrt{\frac{1}{N}\sum x^2}$$

where N is the number of items and x the deviation from the mean, M. The mean deviation from the mean is given by

$$\text{m. d.}=\frac{1}{N}\sum |x|.$$

At times, the mean deviation from the median rather than from the mean is used. The third measure, the quartile deviation Q, is defined by the equation

$$Q=\frac{Q_3-Q_1}{2}$$

where Q_1 and Q_3 are the first and third quartiles. A fourth measure, the range between the highest and lowest recorded values, is also used as a measure of variability.

These measures of variation are useful in helping to understand the precipitation regime in any one locality. However, they are not satisfactory in helping to compare the variations in two different localities. For instance, where the mean annual rainfall is 70 inches, a drop of 10 inches from one year to the next means little to most people; where the mean annual fall is 12 inches, a difference of 10 inches may mean the difference between abundance and drought. Relative measures of precipitation variability have been computed in order to derive comparable figures. There are four of these, corresponding to the four measures of variability. They are:

the coefficient of variation, $v_\sigma=\frac{\sigma}{M}$

the relative variability, $v_r=\frac{m.\ d.}{M}$

$$v_Q=\frac{Q}{Median}$$

and $\frac{Range}{Median}$.

A fifth measure, the "Variability Index" was developed by Maurer [9,10] (also see Loewe [8]). If r is the rainfall and s the variability index, the two are connected through the equation

$$r=a\,(b^s-1)$$

or

$$s=\frac{1}{\log b}\log\left(\frac{r}{a}+1\right).$$

In these equations, a and b are constants for which Maurer finds empirically the values of 6.476 and 1.18 respectively when r is in inches. According to Maurer, this measure does not give an exceptionally high value in those instances where the total rainfall is low and so where the variability has no practical significance. The measure using the range of precipitation was used by Liu En-Lan [7] in his discussion of precipitation over China but then discarded in favor of v_Q. Biel [2] used the second of these four, and computed the relative variability for a large number of world stations. This measure has been used considerably by other climatologists. Conrad [4] and Conrad and Pollack [5] discussed these values and questioned their use as a comparative measure of variability.

The need for some standard of variability for Canadian stations demanded a study of these various measures to determine if any is suitable and if so, which one is the best. This is a report on this study.

COMPARISON OF MEASURES OF VARIABILITY

In discussion of statistical measures Yule and Kendall [12] list five desiderata. A measure should be:

(1) Rigidly defined
(2) Based on all observations
(3) Readily comprehensible
(4) Calculated with reasonable ease
(5) Little affected by fluctuations in sampling.

Of the five measures of variability, the one using the extremes is the only one not satisfying condition (2). For this and other reasons, it is seldom considered. Maurer's Variability Index fails to satisfy the requirement (3).

The other three measures all satisfy condition (1), and in varying degrees conditions (3) and (4). The degrees to which the three measures satisfy condition (5) have never been fully answered. With some statistical measures, the reliability can be determined by computing the probable error. In the present problem, this is impossible. The formulae for probable error are based upon the hypothesis that the distribution approximates the normal curve of error. When the precipitation is large, this is approximately true. When the total is small, the distribution approaches the incomplete gamma function, according to Barger and Thom [1], and the formulae for the probable errors are not valid. Yet the variability is of interest in such distributions.

The use of the formulae for probable error also assumes that the sampling is random. With precipitation values, this is not generally true. In most instances, the distribution includes all available data, which consists of precipitation values for a series of consecutive years. Because precipitation values often appear to be cyclical in nature, there is some correlation among the selected values. This destroys the randomness of the selection.

For the reasons given above, it was necessary to devise some other method of determining the degree to which the three measures, v_σ, v_r, and v_Q, are affected by our "sampling" methods of taking the available data. Several long series were selected. In each series, the values of v_σ, v_r, and v_Q were determined for 10, 15, 20, 25 . . . years for as long as the series continues. Halifax, Nova Scotia, and Qu'Appelle, Saskatchewan, were two of the selected stations. Halifax represents those stations which have a small variation and Qu'Appelle those with large variations. The series of monthly values were used as well as the annual values, because the variations would be larger. Annual data for twelve Chinese stations were treated the same way. These data were taken from Chu [3], and were those used by Liu En-Lan in his study of variation in Chinese rainfall. Figure 1 gives the change in the values of the different measures of variation as the series lengthens. Only a few of those calculated are shown. The values are plotted on semi-logarithmic paper in order that the relative variations, not the actual variations are emphasized.

One fact becomes apparent from the diagrams. For short records, there are wide random variations in the values of all three variables. Taking these results as a guide, one concludes that in regions where the fluctuations are large, it is necessary to have a distribution of at least 30 items before one can be confident that a value is not affected too greatly by the small number in the sample. This figure is approximately the same as that suggested by Landsberg and Jacobs [6] for precipitation normals. They state that for precipitation data, a period of 25 years is necessary for island stations before one can have confidence in the mean values. A 30-year period is suggested for shore stations, 40 for plains, and 50 for mountainous districts.

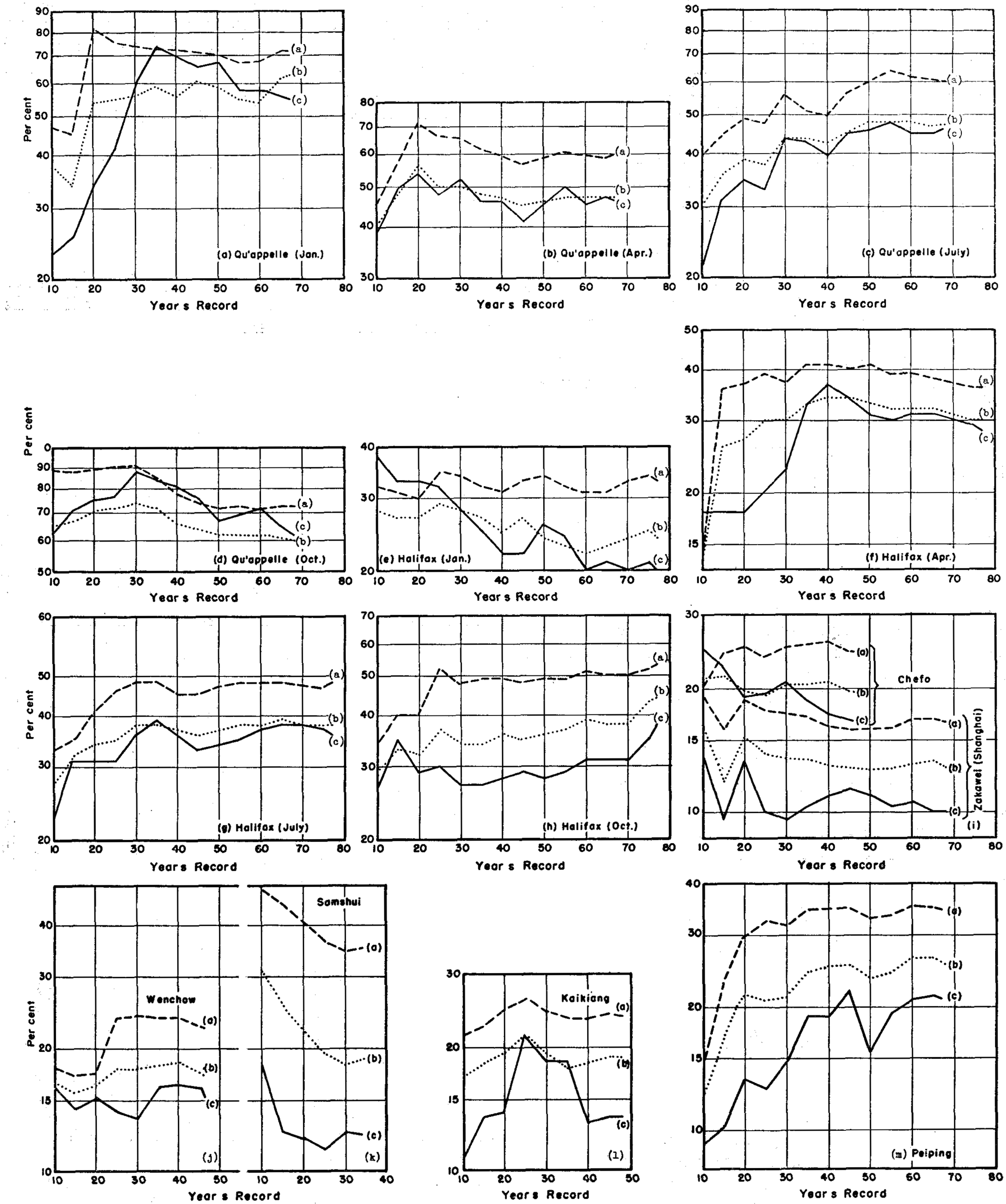

FIGURE 1.—The variation with increasing length of period of three measures of variability of precipitation at selected Canadian and Chinese stations.
(a) coefficient of variation, σ/M;
(b) relative variability, m. d.$/M$;
(c) interquartile variability, Q/Median.

Even a 30-year series is not long enough to make us confident that no major changes will occur with a lengthening record. For instance, for October at Qu'-Appelle, the value of v_σ drops by about 20 percentage points as the series increases in length from 30 to 50 years.

The second conclusion is that there is no major difference between the changes in value of v_σ and v_r, but that the value of v_Q fluctuates much more after there are more than 30 years' data. Evidently v_Q is more affected by variations in sampling than either of the other measures.

The chart showing the values of Samshui is included in figure 1 because it illustrates a fact emphasized by Liu En-Lan. Here the value of v_Q changes less after 15 years than either of the other measures. The third item in the series for Samshui is 5,322 mm., when the next largest value is 2,401 mm., and the mean value 1,784 mm. The extreme value has a greater influence on the values of v_σ and v_r than on v_Q, but its influence on the first two measures gradually decreases as the period of observations becomes longer. Because an extreme value has less influence on the values of v_Q, Liu En-Lan chose this as the best measure of variation for all series. But as can be seen in figure 1, this is true only in extreme instances, and in general, the values of v_Q are more affected by the sampling methods in collecting precipitation statistics.

Of the two measures, the coefficient of variation and the relative variability, figure 1 shows very little difference. Referring back to the desiderata, it will be noted that there is little to choose between the two. Because it can be manipulated mathematically and because in distributions which approximate the normal curve it is particularly useful, statisticians prefer to use the standard deviation as a measure of variation unless there is some very definite reason for preferring another measure. (See Yule and Kendall, [12], p. 144.) The standard deviation and the coefficient of variation were selected as measures of the variation of rainfall for Canadian stations.

BEHAVIOR OF v_σ AND v_r FOR SMALL MEANS

The close relationship between v_σ and v_r in figure 1 is recognized by statisticians. In general, with normal and moderately skew distributions,

$$m.\,d.=\frac{4}{5}\,\sigma.$$

The following discussion of v_σ can then be considered to apply to some extent to v_r.

Conrad [4] raised an objection to v_r which if valid indicates a decrease in usefulness for the measures of dispersion. According to his work, "In the case of yearly sums below 1,000 mm. (39.4 inches) a very strong mathematical dependence of the value of v_r on the yearly sum occurs." He concludes then that one cannot use the measure to compare the variability of precipitation in a locality with a small annual fall with that in another locality where the precipitation is large. It is well to examine this conclusion further.

The formula for the relative variability is:

$$v_r=\frac{1}{MN}\sum|x|.$$

Also

$$v_\sigma=\frac{1}{M}\sqrt{\frac{1}{N}\sum x^2}$$

The individual items of the series of precipitation are all positive or zero, and so the mean is zero only if all the terms and deviations are zero.

Now if

$$f(a)=g(a)=0$$

then

$$\lim_{x\to a}\frac{g(x)}{f(x)}=\frac{g'(a)}{f'(a)}.$$

Since in both v_r and v_σ the denominators and numerators are of the first degree in x, the limit when the mean approaches zero is finite.

TABLE 1.—*Means, standard deviations, and coefficients of variation of precipitation at selected British Columbia and Washington stations*

Station	July			December			Annual			Years record
	Mean (M)	Stand. Dev. (σ)	Coeff. of Var. (v_σ)	Mean (M)	Stand. Dev. (σ)	Coeff. of Var. (v_σ)	Mean (M)	Stand Dev. (σ)	Coeff. of Var. (v_σ)	
British Columbia	(*in.*)	(*in.*)	(%)	(*in.*)	(*in.*)	(%)	(*in.*)	(*in.*)	(%)	
Agassiz	1.96	1.21	62	7.45	3.79	51	58.68	12.70	21.6	57
Alberni	1.26	.99	78	11.77	4.46	38	66.50	12.36	18.6	48
Atlin	1.14	.80	70	1.04	.66	64	10.99	2.70	24.5	31
Barkerville	3.56	1.70	48	3.69	1.70	46	40.21	9.37	23.3	56
Bella Coola	2.07	1.11	54	7.11	3.58	50	54.90	11.90	21.7	45
Big Creek	1.46	.95	65	1.14	.90	79	12.27	2.11	17.2	42
Clayoquot	2.79	2.06	74	16.08	5.78	36	106.30	22.24	20.9	41
Fort St. James	1.67	.90	54	1.50	.65	44	15.62	3.02	19.3	55
Garry Point	.96	.91	95	5.72	2.00	35	36.46	5.87	16.1	48
Glacier	2.67	1.39	52	7.68	3.32	43	56.90	8.39	14.8	38
Golden	1.31	.88	67	2.29	1.72	75	18.31	3.20	17.5	35
Greenwood	1.02	.88	86	1.87	1.06	57	17.04	5.77	33.8	35
Hedley	.91	.82	90	1.14	.73	64	11.45	2.96	25.9	39
Hope	1.40	1.09	78	8.97	3.84	43	56.39	9.72	17.2	36
Kamloops	1.00	.78	78	1.03	.77	75	10.21	2.18	21.3	55
Kelowna	.81	.72	89	1.47	.94	64	12.22	2.40	19.6	46
Masset	3.00	1.76	59	6.98	2.54	36	55.17	9.38	17.0	46
Nanaimo	.99	.85	86	6.53	2.62	40	37.91	7.75	20.4	49
Nelson	1.49	1.17	79	3.58	1.79	50	26.32	5.34	20.3	44
New Westminster	1.35	1.14	84	8.41	3.18	38	57.06	7.47	13.1	68
Prince George	2.11	1.12	53	2.03	1.05	52	22.31	4.61	20.6	33
Prince Rupert	4.81	2.93	61	11.08	4.53	41	93.88	13.54	14.4	36
Princeton	.97	.82	85	1.83	1.20	66	13.17	3.01	22.9	53
Quatsino	2.24	1.61	72	15.20	5.25	35	92.74	18.55	20.0	46
Quesnel	2.00	1.12	56	1.46	.90	62	16.88	3.46	20.5	46
Revelstoke	2.06	1.48	72	5.37	2.03	38	39.43	6.44	16.4	32
Rossland	1.16	.89	77	3.62	1.28	35	29.08	5.32	18.3	38
Vancouver	1.35	1.12	83	8.31	2.86	34	58.02	8.12	14.0	42
Vernon	1.20	.98	82	1.62	1.01	62	14.80	2.70	18.2	48
Victoria	.47	.42	89	5.36	2.89	54	28.68	7.01	24.4	73
Washington										
North Head	.88	.70	80	8.85	3.24	37	54.38	11.18	20.6	51
Seattle	.59	.52	88	5.19	2.47	48	32.15	6.21	19.3	51
Spokane	.41	.52	127	1.98	.82	41	15.05	2.87	19.1	51
Tatoosh Island	1.70	1.51	89	11.63	3.18	27	30.64	11.08	36.2	48

Schumann and Mostert [11] showed that there is an upper limit of 200 percent for v_r. When the series consists of $N-1$ items with value zero, and one other item of any value whatsoever, then

$$v_r = 2 - \frac{2}{N} \tag{1}$$

In a manner similar to that of Schumann and Mostert, it can be shown that v_σ has its maximum value with such a series, its value being

$$v_\sigma = \sqrt{N-1}. \tag{2}$$

Something of the nature of these two quantities given by equations (1) and (2) can be seen by considering the variability of daily precipitation. If the series of daily amounts of precipitation begins with the beginning of a dry spell, the first items are all zero. Both v_r and v_σ are zero until the first rainy day. Then v_σ and v_r take on positive values dependent upon the length of the dry spell, but not dependent upon the amount of rain which fell. If now there is another period without rain, both increase but v_σ increases more rapidly. With the inclusion of another rainy day into the series, the different amounts of rain add other variables to the problem. However, usually, the second rainy day reduces the variability, whether it is measured by v_σ or v_r.

From the above example, we see that both measures remain finite with small values of the mean, but that the coefficient of variability is more sensitive to changes under these conditions than is the relative variability. In the extreme case, where there is only one item other than zero, the value of the variability is independent of the size of the mean.

RELATIONSHIP BETWEEN VARIABILITY AND MEAN PRECIPITATION

In the light of the preceding discussion, one is permitted to wonder whether variability is a function of mean precipitation or not. Is the difference in the measure of variability for a station in Egypt and one in England a result of the difference in total precipitation? Or are those climatic factors which cause rain to fall more erratic in Egypt than in England? Or are both necessary as an explanation of the difference between the two?

The province of British Columbia is an excellent region in which to test the relationship between precipitation and variability. The annual precipitation varies from over 150 inches at some coastal stations to less than 10 inches in some interior valleys. The topography is the cause for the spatial variation; but the variation from year to year is determined by the strength and location of the current which brings the moist Pacific air over the province.

Thirty British Columbia stations with records of at least 30 years were found. For each of these stations, the series of July, December, and annual precipitations were chosen. Table 1 gives the mean, standard deviation, and coefficient of variation for each of these series. Also included are data from four stations in the neighboring state of Washington. The months of July and December were selected as representing the months of greatest and least precipitation and probably of the least variation for the province. The mean values given in the table were computed from grouped data, and so are not necessarily identical with the means usually published for these stations.

The values of the coefficient of variation are plotted on the maps given in figures 2, 3, and 4 and then isolines of coefficient of variability are drawn. In drawing the isolines, the author has used the available data, but recognizes that further information could cause radical changes in these lines.

Figure 2 gives the variability of annual precipitation. In most areas, the variability is near or under 20 percent. Only in two areas is the variability much greater. One is along the Juan de Fuca Strait and includes Victoria

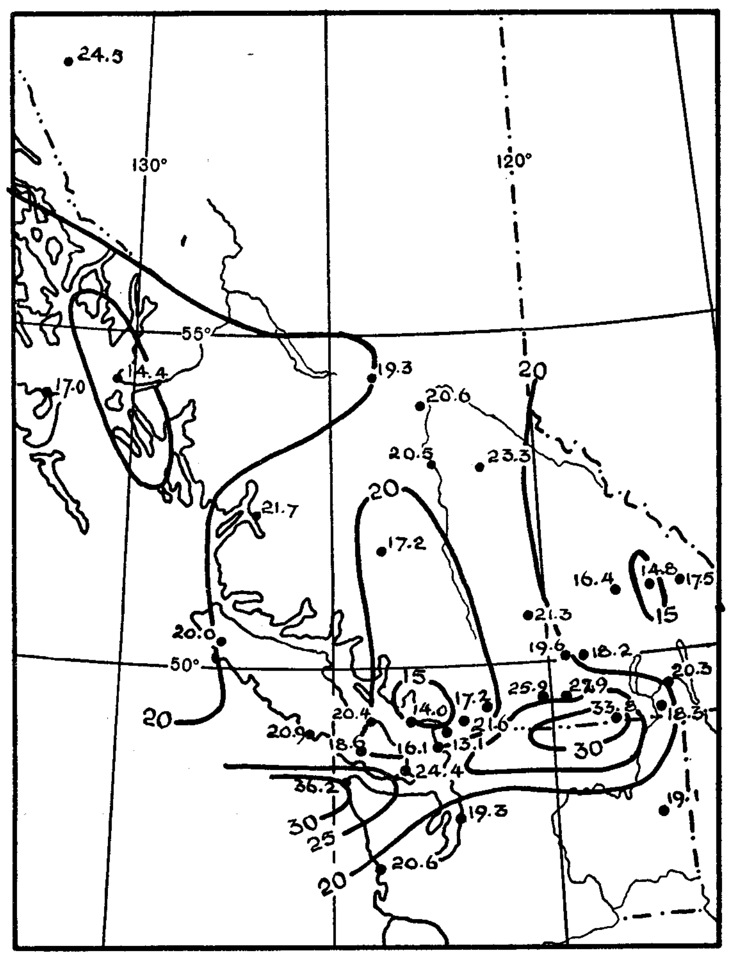

FIGURE 2.—Coefficient of variation (v_σ) of annual precipitation, British Columbia.

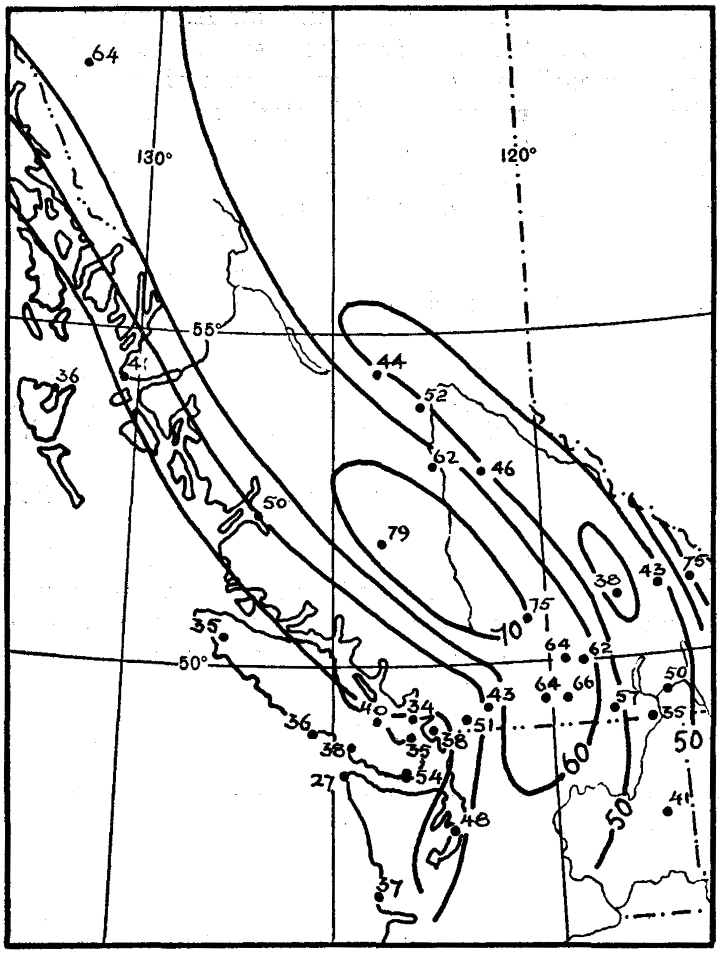

FIGURE 3.—Coefficient of variation (v_σ) of December precipitation, British Columbia.

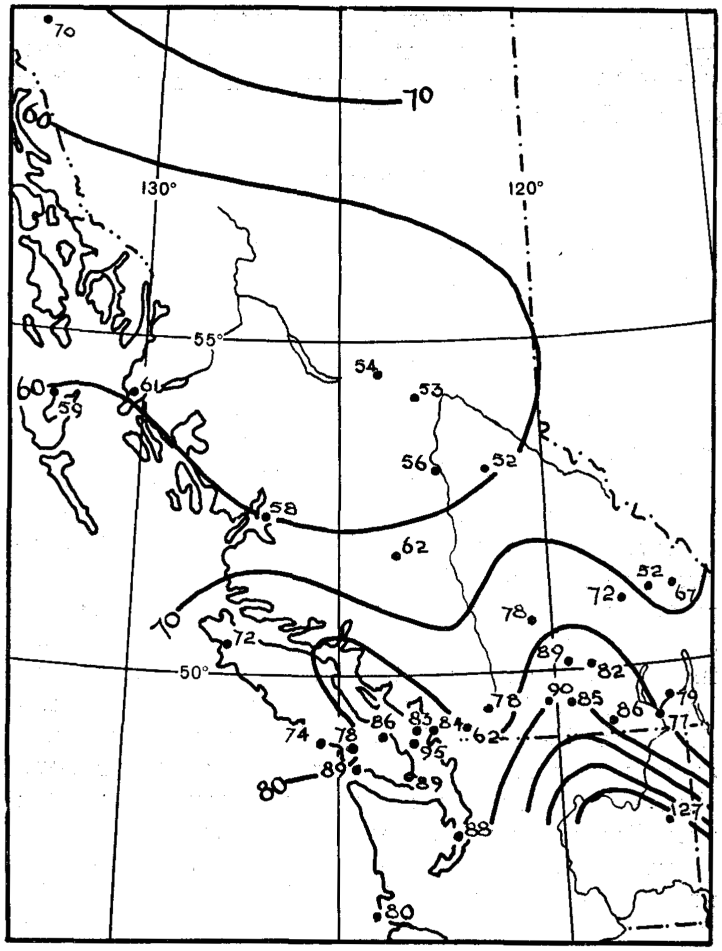

FIGURE 4.—Coefficient of variation (v_σ) of July precipitation, British Columbia.

and Tatoosh Island. The other includes the stations of Greenwood in the Columbia River Valley and Hedley in the Okanagan Valley. The regions of least variability are found around Prince Rupert, at the mouth of the Fraser River, and near Glacier in the Selkirks. In all three regions, the effect of the flow of air from the Pacific is relatively unobstructed by the mountain ranges to the west. But there must be other factors since Clayoquot, Bella Coola and Tatoosh all have greater values of v_σ although they are in the direct path of the westerly winds.

The map for December (fig. 3) shows a much smoother pattern for the values of v_σ. Minimum values are found along the coast and along a parallel line near the Rockies extending from Fort St. James to Rossland. Maximum values exist just east of the Coast Range and at Golden in the Rocky Mountain Trench.

The most reliable July precipitation (see fig. 4) is found in the central section of British Columbia where storms move off the Pacific most regularly. The variability increases as one moves south with Spokane having the greatest value found for v_σ, a value of 127 percent.

A study of the maps and the values given in the table gives evidence that the variability tends to be greater where the precipitation is least. However, the relationship is not close. The coefficient of correlation between mean precipitation and coefficient of variability for July data is −0.68, for December data −0.71, and for the annual data −0.48. The patterns of isolines as drawn on the maps indicate that the values of v_σ are comparable even when the mean precipitation values are quite different. On the basis of this study, one can then conclude that the coefficient of variation is a satisfactory measure for comparing the variability of precipitation between two stations.

SUMMARY

Of the measures of variability, the coefficient of variation was found to have small errors through sampling, and is satisfactory for comparing variability between different stations. It is then considered the most satisfactory measure of variability of precipitation.

ACKNOWLEDGMENTS

The author wishes to thank Mr. C. C. Boughner, Superintendent of Climatology, for his help and interest in the study, and Mr. Andrew Thomson, Controller of the Meteorological Service of Canada, for permission to publish this report.

REFERENCES

1. G. L. Barger and H. C. S. Thom, "Evaluation of Drought Hazard," *Agronomy Journal*, vol. 41, No. 11, November 1949, pp. 519–526.
2. E. Biel, "Die Veränderlichkeit des Jahressumme des Niederschlags auf der Erde," *Geographische Jahresberichte aus Österreich*, Bd. XIV & XV, Leipzig 1929, pp. 151–180.
3. Coching Chu, Chang Wang Tu and Pao-Kun Chang, *The Chinese Rainfall*, Nanking 1935.
4. V. Conrad, "The Variability of Precipitation," *Monthly Weather Review*, vol. 69, No. 1, January 1941, pp. 5–11.
5. V. Conrad and L. W. Pollack, *Methods in Climatology*, Harvard University Press, Cambridge, Mass., 1950, pp. 54–56.
6. H. E. Landsberg and W. C. Jacobs, "Applied Climatology," *Compendium of Meteorology*, American Meterological Society, 1951, pp. 976–992. (See p. 979.)
7. E. L. Liu, "Rainfall Variability or Reliability," *Journal of the Chinese Geophysical Society*, vol. 2, 1950, pp. 191–215.
8. F. Loewe, "Some Considerations Regarding the Variability of Annual Rainfall in Australia," *Bulletin 39*, Australian Meteorological Branch, 1948, pp. 1–13.
9. H. Maurer, "Zur Methodik der Untersuchungen über Schwankungen der Niederschlagsmengen," *Meteorologische Zeitschrift*, vol. 28, 1911, pp. 97–114.
10. H. Maurer, "Das Schwankungsmass der Jährlichen Niederschlagsmengen," *Meteorologische Zeitschrift*, vol. 45, 1928, pp. 166–174.
11. T. E. W. Schumann and J. S. Mostert, "On the Variability and Reliability of Precipitation," *Bulletin of the American Meteorological Society*, vol. 30, No. 3, March 1949, pp. 110–113.
12. G. U. Yule and M. G. Kendall, *An Introduction to the Theory of Statistics*, 13th edition, Charles Griffin and Co., Ltd. London, 1946, pp. 113–144.